Volume 12 • Number 1 • 2009
VALUE IN HEALTH

# The Development of the Herpes Symptom Checklist and the Herpes Outbreak Impact Questionnaire

Lynda C. Doward, MRes,[1] Stephen P. McKenna, PhD,[1,2] David M. Meads, MSc,[1] Kristijan Kahler, RPh, SM,[3] Feride Frech, MPH[3]

[1]Galen Research, Manchester, UK; [2]School of Psychology, University of Central Lancashire, Preston, UK; [3]Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA

## ABSTRACT

**Objective:** To develop measures of the day-to-day symptomatic and functional impact of recurrent genital herpes (RGH) outbreaks. The Herpes Outbreak Impact Questionnaire (HOIQ) and the Herpes Symptom Checklist (HSC) were designed to be acceptable to clinical professionals and to reflect patients' experience.

**Methods:** Scale content was derived via literature review and interviews with RGH patients and physicians. Questionnaires were assessed for face/content validity in the UK and the language checked for acceptability in the United States. The US measures were assessed for face/content validity with patients. Scaling/psychometric properties were determined via web survey. Participants completed the questionnaires twice during an outbreak, with 24 to 72 hours between administrations.

**Results:** Respondents found the questionnaires relevant and easy to understand and complete. Application of Rasch analysis resulted in the removal of two HOIQ items. Both scales were found to be unidimensional. Item stability testing for the HOIQ indicated that the measure is reproducible. Internal consistency was good (alpha: time 1 = 0.87, time 2 = 0.91). Discriminative validity was demonstrated by the measure's ability to distinguish between individuals who differed by self-reported severity of outbreak. The HOIQ and HSC were both responsive to change over time.

**Conclusions:** The HSC and the HOIQ can determine the impact of a herpes outbreak effectively. They are designed to be used daily during such outbreaks and to determine the effectiveness of RGH treatment.

*Keywords:* functioning, patient reported outcome, questionnaire, recurrent genital herpes, symptoms.

## Introduction

Genital herpes is one of the most common viral sexually transmitted diseases, affecting millions of people worldwide [1]. It is caused by infection with the herpes simplex virus HSV-1 (the form more commonly associated with orolabial lesions) or HSV-2 (traditionally associated with sexually acquired anogenital lesions) [2,3]. The condition is thought to be prevalent in all sexually active populations, with rates rising in adolescents with a peak in 20- to 29-year olds [4,5].

For many infected individuals, the condition is benign and self-limiting; up to three quarters experience no, or very mild, symptoms and may be unaware that they have the disease [6]. Nevertheless, the remainder experience recurrent outbreaks of symptoms that can vary in severity and duration. Patients experiencing recurrent genital herpes (RGH) experience an average of four to six episodes a year, although up to one-third of patients report in excess of 10 episodes [7,8]. Recurrent episodes are typically characterized by a period of prodromal symptoms (itching, tingling, inflammation and localized pain along with general flu-like symptoms) followed by emergence of genital sores or ulcers that may take up to six days to heal [8,9]. Frequency and severity of recurrences may diminish naturally over time and current treatments can result in a reduction in the number or severity of outbreaks or reduce viral shedding. Nevertheless, no cure exists and many individuals continue to experience distressing and painful outbreaks for much of their adult life [10].

*Address correspondence to:* Lynda C Doward, Galen Research, Enterprise House, Manchester Science Park, Lloyd Street North, Manchester M15 6SE, UK. E-mail: ldoward@galen-research.com

10.1111/j.1524-4733.2008.00424.x

The impact of RGH on the quality of life (QoL) of affected individuals has been well documented [8,11]. Uncertainty surrounding frequency and severity of outbreaks and the perceived stigma attached to having a sexually acquired infection has a profound impact on personal relationships, feelings of sexual desirability, sexual activity, mood, well-being and, ultimately, self-esteem [11,12]. Fear of forming relationships and of rejection may affect the capacity for expressing physical warmth and intimacy and the enjoyment of sex. The disorder may also lead to a loss of sexual confidence and adversely affect feelings of sexual desirability [13].

Although many studies have focused on the long-term psychological impact of RGH, limited information is available on the experience of herpes outbreaks and no measures specific to day-to-day impact exist. The Recurrent Genital Herpes QoL Questionnaire (RGHQoL) [14] has been shown to be highly responsive to changes in QoL associated with antiviral treatment [7]. Nevertheless, this measure is designed to be used between outbreaks and is concerned with longer-term issues related to the impact of the disease and its treatment on the patient. During herpes outbreaks, the patient is likely to be concerned with more immediate impacts, such as pain or restrictions in social and work activities. Such factors are related to impairment and functioning rather than QoL.

The purpose of this study was to develop instruments to measure day-to-day symptomatology and functioning that are specific to genital herpes outbreaks. Two complementary instruments were developed; the Herpes Symptom Checklist (HSC) to assess daily symptoms and the Herpes Outbreak Impact Questionnaire (HOIQ) to measure day-to-day functional impact of outbreaks. Both instruments were designed to:

 

1 Be easy to administer;
2 Be acceptable to clinical professionals;
3 Reflect the experience of RGH patients;
4 Have good psychometric properties (specifically; unidimensionality, reliability and validity).

Because the instruments were initially required to establish the effectiveness of antiviral therapies during the course of an outbreak, there was a specific need for the scales to be responsive to change in impact over a short period of time. The measures are intended to be used during actual outbreaks only.

## Methods

### *Study Participants*

Patients involved in all stages of the research were required to meet the following inclusion/exclusion criteria:

*Inclusion criteria*

1 18 years of age or older;
2 Established RGH;
3 One or more of the following should also apply:
   - Experiencing three or more herpes outbreaks a year, with RGH duration of 12 months or more.
   - If experiencing fewer than three outbreaks per year, patients should have had RGH for at least 12 months and currently be on suppression treatment, providing that they had experienced an outbreak frequency of three or more occurrences a year before suppression.
   - If they had RGH for less than a year, patients must have had herpes for at least 6 months and have experienced more than three outbreaks in that time.

*Exclusion criteria.* Patients with any major co-morbidity considered likely to influence their responses to the questionnaires.

### *Generation of Content for the New Questionnaires*

The methodology employed was designed to generate questionnaire items while checking that issues considered important by clinical experts were not omitted from the final measures. Three methods were used to generate questionnaire content; interviews with patients, a literature review and interviews with clinical experts.

*Interviews with RGH patients.* Interviews were conducted with 25 RGH patients recruited through Genitourinary Departments in the UK. Emphasis during the interviews was placed on the actual herpes outbreak with respondents encouraged to describe in detail the symptoms they experienced and the way in which their social, emotional and physical functioning was impaired. The interviews were audio-recorded and full transcripts produced.

To supplement these interviews, the researchers had access to 15 transcripts of interviews conducted with RGH patients as part of the development of the RGHQoL during the early 1990s [14]. These interviews included patient reports of symptoms and functioning arising from herpes outbreaks.

The transcripts from both sets of interviews were subjected to thematic analysis to identify key areas of impact. Analysis was conducted independently by three experienced researchers and their findings compared. Issues related to symptoms or functioning were identified and recorded.

*Review of the literature.* A literature review was undertaken to identify any additional information on the symptomatic or functional impact of RGH. Findings from the review were compared with the results of the analysis of patient interviews to ensure that no areas had been missed.

*Interviews with clinical experts.* Interviews were conducted with three clinical experts employed in Genitourinary Departments in the UK who specialized in treating sexually transmitted diseases. They were asked to review the list of symptoms and functional problems generated after the analysis of patient interviews and review of the literature to identify whether any issues of importance had been omitted.

Draft measures were designed using, as far as possible, the wording/phrasing used by interviewees.

### *Cognitive Debriefing Interviews*

Interviews were conducted with 12 patients in the UK to assess the face and content validity of the draft questionnaires. Interviewees were asked to complete the questionnaire in the presence of the interviewer and then to comment on the suitability and acceptability of the content. They were also asked to indicate whether any important issues had been omitted.

### *Production of US-English Versions of Scales*

The first requirement for use of the measures was in the United States. To ensure that the instruments were appropriate for use in the United States, the UK content was adapted to US English by a lay panel held in the United States. Emphasis was placed on ensuring that the wording selected was appropriate for the local population. Cognitive debriefing interviews were then conducted with nine RGH patients in North Carolina to ensure the face and content validity of the US measures using the same methodology as above.

The following stages of the study tested the psychometric properties of the US version of the scales, including unidimensionality, reliability and construct validity.

### *Assessment of Scaling and Traditional Psychometric Properties*

A web survey was conducted to assess the scaling and psychometric properties of the draft scales via the Antopia Web site, http://www.MPwH.net. Antopia is an Internet-based organization providing social resources, support and information for people living with herpes.

Both the HSC and HOIQ were completed together with a brief demographic questionnaire. The latter also included questions on RGH history and perceived general well-being and severity of RGH.

Only members of MPwH who paid for upgraded services were able to complete the survey. Interested respondents were required to complete screening questions to ensure that they met the eligibility criteria before completing the HSC and HOIQ. An automated online program was devised to ensure that eligible subscribers could complete only one set of responses.

Potential responders were asked to complete the questionnaires on two occasions: on the first day of a herpes outbreak and again 2 days later. Respondents who completed the second set of questionnaires fewer than 24 or more than 72 hours after completing the first were excluded from the study.

*Scaling assessment.* Rasch analyses were conducted on the survey data to assess the scaling properties of the HOIQ. The

**Table 1** Characteristics of study participants

| | UK Interview sample | UK field-test interviews | US field-test interviews | US Web survey |
|---|---|---|---|---|
| Sample size | 25 | 12 | 9 | 158 |
| Number (%); men | 8 (32) | 3 (25) | 4 (44.4) | 76 (47.1) |
| Number (%); women | 17 (68) | 9 (75) | 5 (55.6) | 82 (51.9) |
| Mean (SD) age; years | 40.9 (10.1) | 40.5 (8.0) | 39.0 (7.5) | 35.8 (7.6) |
| Age range; years | 18.0–65.0 | 31.0–55.0 | 27.0–49.0 | 19.0–59.0 |
| Number (%) employed | 18 (72) | 9 (75) | | 137 (86.7) |
| Mean (SD) duration of herpes; years | 11.9 (6.2) | 12.7 (5.1) | 6.9 (7.1) | 8.6 (7.3) |
| Duration range; years | 2.5–23.0 | 5.5–20.0 | 0.33–21.0 | 0.5–30.0 |
| Mean (SD) number of outbreaks per year | 7.8 (6.8) | 6.2 (3.3) | | 5.3 (3.2) |
| Mean (SD) duration of outbreaks; days | 8.4 (4.6) | 7.1 (4.8) | | 6.4 (3.9) |

SD, standard deviation.

Rasch model [15] utilized is the one-parameter logistic item response theory model that is most widely used in health outcome measurement. Analyses were conducted using the Rasch Unidimensional Measurement Model (RUMM2010) software [16] adopting the unrestricted partial credit model.

The adequacy of the fit to the model was evaluated through a total chi-square ($\chi^2$) fit statistic. In addition, the fit of the individual items was evaluated through individual item $\chi^2$ fit statistics. A statistically significant $\chi^2$ is taken to indicate inadequate fit to the model. Given the number of tests applied, a 95% significance level would not normally be considered appropriate. Thus, fit values below 0.01 were taken to indicate misfit and those between 0.01 and 0.05 were taken as indicating borderline misfit. Further information on the consistency of the pattern of responses to individual items was provided by standardized residuals ([observed score – expected score]/standard error). Residuals within the range 2.5 to –2.5 were taken as indicating adequate consistency. Items that significantly misfit or had excessive residuals were considered candidates for removal in the item reduction process.

The logit and person coverage of the HOIQ was examined using the item map of uncentralized thresholds and the person–location distribution map. This informs on the scope of the underlying construct covered by the scales.

Rasch analysis was also used to assess unidimensionality of the HSC. Confirmation of unidimensionality allows item scores to be summed to produce a single-scale score.

*Traditional psychometric analyses.* Having identified the final item set for the HOIQ, traditional psychometric tests were applied to the survey data. These included assessment of internal consistency (Cronbach's alpha coefficients) and construct validity.

The assessment of test–retest reliability was not possible because of the rapid changes that occur with the healing of lesions. Instead, item stability was assessed by examining differential item functioning (DIF; see [17,18] for more details) within the Rasch framework. Survey data were analysed to determine whether DIF occurred by questionnaire administration time. For a measure to be reliable, the ordering of items (in terms of severity) should be the same each time the questionnaire is administered and there should be no DIF according to administration.

In addition to fit to the Rasch model, construct validity was evaluated by assessing the level of association between scores on the HOIQ and those from the HSC. HOIQ scores were correlated with HSC scores using nonparametric Spearman rank correlation coefficients. The HOIQ was expected to show moderate association with the HSC because they measure related but distinct constructs.

Discriminative validity was assessed by testing the ability of the measures to distinguish between groups of individuals that differed according to some known factor. The factors used for both scales were: patient-perceived severity of current and typical outbreak (mild, moderate, quite severe or very severe) and average duration of outbreaks (days). Non-parametric tests for independent samples (Mann–Whitney *U*-test for two groups or Kruskal–Wallis One-way Analysis of Variance for three or more groups) were employed to test for differences in scale scores between groups.

Because responsiveness was a crucial requirement of the measures, it was necessary to show that they were capable of demonstrating improvements with time since start of outbreak. It was hypothesized that scores on both measures would be lower (indicating an improvement in symptomatic and functional impact) on day 3 than day 1.

## Results

### *Sample Details*

Table 1 shows demographic characteristics of participants at each stage of the study. Information on mean number and duration of outbreaks was not collected for US field-test participants. Nevertheless, all had experienced an outbreak within the previous 2 months. Participants who completed the Internet-based survey also provided information on the perceived severity of their current and typical outbreaks (Table 2).

### *Generation of Questionnaire Content*

The following activities were affected or avoided:

- Social activities. Respondents reported avoiding places or activities where they might become too hot during outbreaks.
- Sporting activities. For example, interviewees would return home to shower rather than using facilities at the sports center.
- Work. Some interviewees reported being unable to attend work because of difficulty sitting, feeling generally ill or because of a constant urge to run to the restroom to scratch.

**Table 2** Severity of outbreaks for Web survey participants (n = 158)

| | Severity of current outbreak (%) | Severity of typical outbreak (%) |
|---|---|---|
| Mild | 98 (62.0) | 76 (48.1) |
| Moderate | 56 (35.4) | 72 (45.6) |
| Quite severe | 4 (2.5) | 9 (5.7) |
| Very severe | 0 (0.0) | 1 (0.6) |

**Table 3** Item fit statistics for the Herpes Symptom Checklist

| Item | Location | Standard error | Residual | $\chi^2$ | Probability |
|---|---|---|---|---|---|
| 1 | –0.81 | 0.10 | 2.21 | 11.27 | 0.01 |
| 2 | –0.40 | 0.10 | –1.01 | 1.05 | 0.79 |
| 3 | –0.53 | 0.10 | –0.73 | 7.84 | 0.05 |
| 4 | –0.51 | 0.09 | –1.55 | 4.04 | 0.26 |
| 5 | –1.10 | 0.15 | 0.68 | 0.68 | 0.88 |
| 6 | –0.98 | 0.09 | –2.07 | 4.52 | 0.21 |
| 7 | 0.91 | 0.16 | 0.49 | 0.56 | 0.91 |
| 8 | 1.03 | 0.25 | 0.88 | 4.69 | 0.20 |
| 9 | 0.85 | 0.17 | –0.89 | 2.85 | 0.42 |
| 10 | –0.13 | 0.10 | –0.30 | 9.15 | 0.03 |
| 11 | –0.86 | 0.09 | 0.57 | 2.53 | 0.47 |
| 12 | 0.55 | 0.15 | 0.18 | 1.38 | 0.71 |
| 13 | 1.97 | 0.28 | –0.50 | 1.95 | 0.58 |

- Intercourse and other intimate contact (including hugging or kissing) that could potentially lead to intercourse.
- Personal hygiene activities that would increase discomfort or pain—e.g., use of soap was kept to a minimum.
- Some patients washed more frequently because they felt unclean.
- Interviewees reported delaying urinating until absolutely necessary.
- Avoidance of clothing that caused friction. Some respondents chose not to wear underwear during an outbreak or avoided tight jeans or trousers.
- Avoidance of foods (including chocolate and garlic) perceived to exacerbate outbreaks.
- Avoidance of tiring activities. Respondents reported trying to take things easier, taking more rest and avoiding late nights.

In addition, many interviewees expressed feeling worried or anxious during an outbreak: either because of concern about potentially passing on the condition or to fear of discovery.

The literature review failed to identify additional impairments or disabilities associated with outbreaks of RGH.

After checking the relevance and the frequency with which both symptomatic and functional issues had been raised by interviewees, the areas of impact were presented to three clinical specialists. These individuals supported the findings from the patient interviews.

## *Production of Draft Questionnaires*

Draft versions of the HSC (13 items) and the HOIQ (14 items) were produced. For the HSC, a 4-point severity response scale (none, mild, moderate, and severe) was employed to rate presence and severity of symptoms on the day of completion. The HOIQ also employed a 4-point response scale (not at all, a little, a lot, very much).

Cognitive debriefing interviews conducted with patients in the UK revealed that the questionnaires were well accepted. Respondents reported little difficulty completing the scales and considered the content to be appropriate.

*Production of US versions of the scales.* Most of the language was considered appropriate for use in the United States. Minor changes were made to the HSC. For example, for the symptom *crusts*, it was considered necessary to follow this with the term *scabs* in parentheses for US respondents. Similarly, minor changes were made to several items in the HOIQ, largely to replace Anglicized spelling with Americanized versions (for example, replacing *socialising* with *socializing*). It was also suggested that the last two response options for the HOIQ be changed to "A fair amount" and "A great deal."

Cognitive debriefing interviews conducted with patients in the United States revealed that they were able to understand the questionnaire's instructions and items without any difficulty. The instructions and response scales were easily understood and accepted by all participants and the language employed was considered appropriate for use in the United States.

## *Assessment of Scaling and Traditional Psychometric Properties*

The survey included 158 participants and the average time between administrations of the questionnaires was 2.1 (standard deviation, SD = 0.3) days.

*Scaling assessment.* The 13-item HSC displayed good fit to the Rasch model. The overall $\chi^2$ probability was nonsignificant ($P$-value = 0.072), indicating an adequate fit of the data to the model. The item fit statistic (mean = –0.16; SD = 1.15) and person fit statistic (mean = –0.20; SD = 0.82) also indicated adequate fit. This suggests that the HSC forms a unidimensional scale. Item fit statistics for the HSC are shown in Table 3.

The 14-item HOIQ exhibited significant misfit to the Rasch model (item-trait $\chi^2$ $P$-value < 0.000). Assessment of the threshold ordering was not conclusive because there was a low affirmation of certain response categories as a result of the sample being mild to moderate. Nevertheless, two misfitting items were deleted (one at a time with the analyses re-run after each deletion) and several more were rescored. The rescoring (collapsing of categories) was necessary because an insufficient number of people had selected the severe categories. The remaining 12 items fit the Rasch model (item-trait $\chi^2$ $P$-value = 0.34; item fit statistic mean = 0.17, SD = 1.36; person fit statistic mean = 0.12, SD = 0.88). The final item fit statistics for the HOIQ are shown in Table 4. The item locations indicated that the mildest item (representing the lowest level of underlying construct) is the item "*has your outbreak reduced your interest in sex?*" (location = –2.06) and the most severe item is "*has your outbreak made daily transportation uncomfortable for you?*" (location = 1.71). The item–person threshold distribution (Fig. 1) indicated that the HOIQ items had good coverage of the underlying construct of activity limitation. No items exhibited significant DIF by administration ($P$-value $\leq$ 0.01) confirming the reproducibility (reliability) of the scale.

*Traditional psychometric properties.* The overall score for the 13-item HSC is obtained by summing the scores for the

**Table 4** Final item fit statistics for the Herpes Outbreak Impact Questionnaire

| Item | Location | Standard error | Residual | $\chi^2$ | Probability |
|---|---|---|---|---|---|
| 1 | 0.10 | 0.10 | –0.47 | 3.50 | 0.32 |
| 2 | 0.58 | 0.11 | 0.19 | 0.78 | 0.85 |
| 3 | 0.33 | 0.12 | 0.92 | 3.08 | 0.38 |
| 4 | 1.71 | 0.14 | –1.14 | 2.05 | 0.56 |
| 5 | –0.69 | 0.09 | –0.23 | 1.74 | 0.63 |
| 7 | 0.53 | 0.18 | –0.37 | 2.13 | 0.55 |
| 8 | –0.23 | 0.09 | 0.78 | 2.06 | 0.56 |
| 9 | –0.52 | 0.12 | 2.54 | 8.70 | 0.03 |
| 10 | 0.28 | 0.12 | –1.16 | 2.63 | 0.45 |
| 12 | –0.37 | 0.12 | –1.59 | 8.17 | 0.04 |
| 13 | –2.06 | 0.11 | 0.47 | 3.26 | 0.35 |
| 14 | 0.34 | 0.12 | 0.01 | 0.77 | 0.86 |

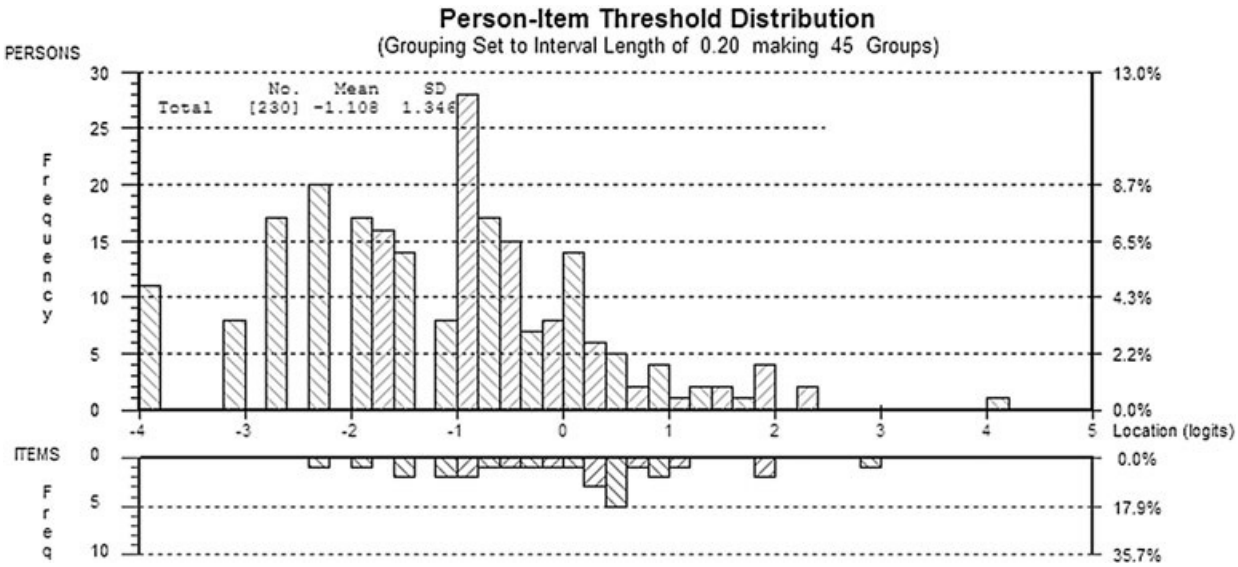


**Figure 1** Item–person threshold distribution for the Herpes Outbreak Impact Questionnaire.

individual symptoms (0–3), giving a possible maximum score of 39. The total score for the 12-item HOIQ is also obtained by summing the scores for the individual items (0–3), giving a maximum potential score of 36. In both cases, a high score indicates greater impact.

Table 5 shows the score statistics for the instruments at each time point. There was a significant decrease in score over time for both the HSC (effect size = 0.50) and the HOIQ (effect size = 0.29), suggest ing that the scales are sensitive to changes that occur during the genital herpes outbreak cycle.

The Cronbach alpha statistic for the HSC was 0.83 at time 1 and 0.84 at time 2 and for the HOIQ, 0.87 at time 1 and 0.91 at time 2 (n = 158 at both administrations). This confirms the internal consistency of the scales. Corrected item–total correlations for HSC and HOIQ scale items were all >0.2 and <0.8, suggesting that all items contribute adequately to their respective scales without being redundant.

Correlations between the HSC and the HOIQ provided evidence of convergent validity. Observed correlations were 0.60 at time 1 and 0.67 at time 2, confirming expectations that the two scales measure related but distinct constructs.

Differences in scores were examined for people who differed according to the duration of their typical outbreak and self-reported severity of their current and typical outbreaks (Table 6). Because of the small number of people who self-reported either their current or typical outbreak severity as being *quite or very severe*, these categories were combined with the *moderate* category to make a single "moderate–severe" group.

As expected, it was shown that individuals who generally had outbreaks of shorter duration scored significantly lower on both scales at time 2 but not at time 1, suggesting that lesions had healed quicker.

Both the HSC and HOIQ were shown to discriminate between individuals based on how severe they rated their current or typical herpes outbreaks. The more severe the ratings of the outbreak, the higher the score on the outcome measures.

There were no significant differences in HSC or HOIQ scores at either time point between people who differed according to time with RGH (above and below the median of 7 years), age (above and below the median age of 36 years) or employment status (those who were and were not working). A significant difference was found between scores on the HSC according to sex at time 1 (with women scoring significantly higher than men). Nevertheless, this difference was not present at time 2 for the HSC or at either time point for the HOIQ.

## Discussion

The RGHQoL questionnaire assesses the overall impact of the infection on the individual's QoL, focusing on the long-term impacts of infection and treatment. Nevertheless, during an

**Table 5** Herpes Symptom Checklist (HSC) and Herpes Outbreak Impact Questionnaire (HOIQ) scores at times 1 and 2

| | HSC (n = 158) | | | HOIQ (n = 158) | | |
|---|---|---|---|---|---|---|
| | Time 1 | Time 2 | *P*-value | Time 1 | Time 2 | *P*-value |
| Mean score | 8.7 | 5.8 | 0.000* | 9.9 | 7.8 | 0.000* |
| Standard deviation | 5.83 | 4.98 | | 7.21 | 7.19 | |
| Median score | 7.0 | 4.0 | | 9.0 | 6.0 | |
| Score range | 0.0–34.0 | 0.0–27.0 | | 0.0–36.0 | 0.0–33.0 | |
| % scoring minimum | 0.6 | 6.3 | | 3.8 | 10.1 | |
| % scoring maximum | 0.0 | 0.0 | | 0.0 | 0.0 | |

*Wilcoxon signed ranks test.

**Table 6** Discriminative validity of the Herpes Symptom Checklist (HSC) and the Herpes Outbreak Impact Questionnaire (HOIQ)

| | | HSC | | | | HOIQ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Time 1 | | Time 2 | | Time 1 | | Time 2 | |
| | n | Median (IQR) | *P*-value | Median (IQR) | *P*-value | Median (IQR) | *P*-value | Median (IQR) | *P*-value |
| Duration of outbreaks | | | | | | | | | |
| 5 days (median) or less | 87 | 7.0 (4.0–12.0) | ns | 4.0 (2.0–7.0) | 0.015* | 8.0 (4.0–13.0) | ns | 5.0 (2.0–9.0)* | 0.000* |
| More than 5 days | 71 | 7.0 (5.0–11.0) | | 6.0 (3.0–9.0) | | 10.0 (5.0–13.0) | | 7.0 (4.0–13.0) | |
| Severity of current outbreak | | | | | | | | | |
| Mild | 98 | 6.0 (4.0–8.0) | 0.000* | 4.0 (2.0–7.0) | 0.003* | 6.0 (3.0–12.3) | 0.000* | 5.0 (1.0–9.0) | 0.000* |
| Moderate-to-severe | 60 | 11.0 (8.0–13.0) | | 6.0 (3.3–9.8) | | 11.0 (8.0–17.0) | | 8.0 (5.0–13.0) | |
| Severity of typical outbreak | | | | | | | | | |
| Mild | 76 | 5.0 (4.0–8.0) | 0.000* | 4.0 (2.0–7.0) | 0.003* | 5.5 (3.0–10.0) | 0.000* | 4.5 (1.0–8.0) | 0.000* |
| Moderate-to-severe | 82 | 9.0 (7.0–13.0) | | 5.0 (3.0–10.0) | | 11.0 (7.0–17.0) | | 8.0 (4.0–13.3) | |

*Mann–Whitney U-test.
IQR, interquartile range; ns, not significant.

outbreak a patient is likely to be more concerned with the symptoms of the actual outbreak (such as pain and discomfort) and short-term limitations to social, work or other activities. Such factors are related to impairment and functioning.

The HSC and the HOIQ are the first instruments developed to measure the impact of individual herpes outbreaks and use items representing impairments and disabilities that were of specific importance to RGH patients. The measures were required to be relevant and comprehensible to patients and to be unidimensional, reliable, valid and easy to administer. They were designed to be used alone or in conjunction with each other.

The instruments have been shown to be acceptable, relevant, clear and understandable to respondents in both the UK and the United States. The scaling and psychometric properties of the US version of the questionnaires have been assessed. Item Response Theory (the Rasch model) was employed to determine the unidimensionality of the HOIQ and HSC. After removal of two items, the HOIQ exhibited good fit to the model, indicating that the 12-item version is a unidimensional measure of functioning during a herpes outbreak and that items may be validly summed to obtain an overall score. The HOIQ items have good coverage of the individuals assessed and the underlying construct, functioning. The HSC data were shown to fit the model without the need for deletion of items.

The traditional psychometric testing of the HOIQ revealed that the scale possesses good psychometric properties. Item stability of the HOIQ items indicated that none of the items exhibited time-related DIF, indicating that the HOIQ is reproducible. The measure was also shown to have good internal consistency. Crucially, the HOIQ also displays evidence of construct validity in the form of discriminative validity. At both time points, the HOIQ was able to distinguish between individuals who differed according to how severe they reported their outbreaks (both current and typical) to be. The HSC also showed evidence of such validity.

The significant decline in scale scores over the cycle of the herpes outbreak (even though the administrations were only 2 days apart) also suggests that the HSC and HOIQ will be responsive to changes in the impact of outbreaks over time. Both measures have been included in a clinical trial that will provide additional information on the responsiveness of the measures. Future analyses of HOIQ and HSC data might also explore whether DIF exists according to age and sex.

There were a number of limitations to the study. First, because the study survey was Web-based, it represents a validation of only that mode of questionnaire administration. Nevertheless, there is evidence to suggest that Web-based surveys provide results comparable to other modes of administration [19,20] and the application of the HOIQ and HSC in a subsequent clinical trial showed that the pen and paper versions of the questionnaires were reliable and valid. The mode of data collection for the validation survey may also have introduced sample bias because it is known that the Internet is used by higher-income, better-educated and younger individuals [21]. Nevertheless, analysis of HOIQ and HSC trial data collected from a clinic indicates that the psychometric properties of the questionnaires were comparable across samples. In addition, the Rasch analysis indicated that some of the response option categories should be collapsed; this may have been as a result of the sample being mild and having a tendency to respond in the mild and moderate response categories. Rasch analysis of HOIQ and HSC data from more severe samples is required to determine whether the response options are appropriate.

Although acknowledging these limitations, it is concluded that the development of the HSC and HOIQ was successful and that these new scales will prove to be valuable outcome measures for use in clinical trials.

## Supplementary Materials

Supplementary material for this article can be found at: http://www.ispor.org/publications/value/ViHsupplementary.asp

The authors would like to thank the members and staff of the Antopia Network and staff and patients at the Sheffield Royal Hallshire Hospital (Dr George Kinghorn). We would also like to thank Novartis Pharma for sponsoring the research.

Source of financial support: This research was supported by Novartis Pharmaceuticals Corporation.

## References

1 Malkin JE. Epidemiology of genital herpes simplex virus infection in developed countries. Herpes 2004;11(Suppl. 1):S2A–23A.
2 Wald A, Zeh J, Selke S, et al. Virologic characteristics of subclinical and symptomatic genital herpes infections. N Engl J Med 1995;333:770–5.
3 Wald A, Zeh J, Selke S, et al. Reactivation of genital herpes simplex virus type 2 infection in asymptomatic seropositive persons. N Engl J Med 2000;342:844–50.
4 Bearman PS, Moody J, Stovel K. Chains of affection: the structure of adolescent romantic and sexual networks. Am J Soc 2004; 110:44–91.
5 Doherty IA, Padian NS, Marlow C, et al. Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. J Infect Dis 2005;191(Suppl):S42–54.
6 Roberts CM. Genital herpes in young adults: changing sexual behaviours, epidemiology and management. Herpes 2005;12: 10–4.

7 Patel R, Tyring S, Strand A, et al. Impact of suppressive antiviral therapy on the health related quality of life of patients with recurrent genital herpes infection. Sex Transm Infect 1999;75: 398–402.
8 Patel R, Boselli F, Cairo I, et al. Patients' perspectives on the burden of recurrent genital herpes. Int J STD AIDS 2001;12: 640–5.
9 HVA. Herpes: frequently asked questions. Herpes virus association web site. Available from: http://www.herpes.org.uk. [Accessed November 26, 2007].
10 Brentjens MH, Yeung-Yue KA, Lee PC, et al. Recurrent genital herpes treatments and their impact on quality of life. Pharmacoeconomics 2003;21:853–63.
11 McKenna SP, Doward LC, Price MJ, et al. Psychological consequences of recurrent genital herpes. Antiinfective Drugs Chemother 1998;16(Suppl. 1):S19.
12 Hunt SM, Kovachich P, McKenna SP. Quality of life in genital herpes. Br J Sex Med 1993;20:12–5.
13 Cattoti DN, Clarke P, Catoe KE. Herpes revisited—still a cause for concern. Sex Transm Dis 1993;20:77–80.
14 Doward LC, McKenna SP, Kohlmann T, et al. The international development of the RGHQoL: a quality of life measure for recurrent genital herpes. Qual Life Res 1998;7:143–53.
15 Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Chicago: University of Chicago Press, 1980.
16 Andrich D. Rasch Models for Measurement. London: Sage Publications, 1988.
17 Holland PW, Wainer H, eds. Differential Item Functioning. Hillsdale: Lawrence Erlbaum Associates, 1993.
18 Brodersen J, Meads D, Kreiner S, et al. Methodological aspects of differential item functioning in the Rasch model. J Med Econ 2007;10:309–24.
19 McCabe SE, Couper MP, Cranford JA, Boyd CJ. Comparison of Web and mail surveys for studying secondary consequences associated with substance use: evidence for minimal mode effects. Addict Behav 2006;31:162–8.
20 Beebe TJ, Locke GR, 3rd, Barnes SA, et al. Mixing web and mail methods in a survey of physicians. Health Serv Res 2007;42: 1219–34.
21 Viswanath K, Kreuter MW. Health disparities, communication inequalities, and eHealth. Am J Prev Med 2007;32(5 Suppl.): S131–3.

## Appendix

### *Example HOIQ Items*

The following questions are intended to find out how your current herpes outbreak has affected you during the last 24 hours.

Please check one box only for each question.

| Thinking about the last 24 hours: has your outbreak | Not at all | A little | A fair ount | A great deal |
|---|---|---|---|---|
| Affected your interest in leisure activities? | ❒ | ❒ | ❒ | ❒ |
| Made walking uncomfortable for you? | ❒ | ❒ | ❒ | ❒ |
| Reduced your interest in socializing? | ❒ | ❒ | ❒ | ❒ |
| Affected the ount you washed yourself? | ❒ | ❒ | ❒ | ❒ |

### *Example HSC Items*

Thinking about today. . . . Look down this list and, for each problem, please put a tick in the box according to whether: You have not had this problem (none), you have had this problem in a mild form, you have had this problem in a moderate form, you have had this problem and it was severe.

Have you had any of the following associated with your outbreak today?

| | | None | Mild | Moderate | Severe |
|---|---|---|---|---|---|
| 1 | Tingling | ❒ | ❒ | ❒ | ❒ |
| 2 | Burning | ❒ | ❒ | ❒ | ❒ |
| 3 | Pain | ❒ | ❒ | ❒ | ❒ |
| 4 | Aching | ❒ | ❒ | ❒ | ❒ |
| 5 | Itching | ❒ | ❒ | ❒ | ❒ |